DOCUMENT RESUME

ED 174 629 TM 008 446

AUTHOR Dirlam, David; Byrne, Maureen
TITLE Standardized Developmental Ratings.
PUB DATE Mar 78
NOTE 20p.; Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31, 1978)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Developmental Stages; *Developmental Tasks; *Discourse Analysis; Educational Diagnosis; Elementary Education; Elementary School Teachers; *Freehand Drawing; Language Development; Reliability; Standardized Tests; Story Telling; Student Evaluation; *Teacher Attitudes; *Test Construction

ABSTRACT

The feasibility of standardized assessment of features of children's organization of knowledge and the demonstration of differences in organization between stages of development were investigated. It was possible to standardize developmentally derived instruments, in the same way as empirically derived tests because such evaluations concern the changing organization of knowledge, rather than content. Drawings and stories were collected from school age children and rated by five teachers trained in an analytic procedure based on Lowenfeld's theory for drawing and Moffett's theory for discourse. It was demonstrated that standardized developmental ratings (SDR) could be done with a high degree of reliability. It was concluded that SDR's could assist teachers in determining a child's developmental level and in detecting when the child made the transition to a subsequent stage of development. (RD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED174629

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Standardized Developmental Ratings

David Dirlam and Maureen Byrne

Educational Research and Demonstration Center
State University College of Arts and Science
Plattsburgh, New York

Presented as part of the symposium:

Alternatives to empirically derived standardized tests

David Dirlam, Organizer
David Irvine, Chairman

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. Dirlam

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

American Educational Research Association
(AERA)

March 1978

Toronto

TM 008 444

Standardized Developmental Ratings

David Dirlam and Maureen Byrne*

Educational Research and Demonstration Center
State University College, Plattsburgh, New York

Introduction

We cannot yet say much in a standardized manner about how people at a given stage of development differ in the way their knowledge is organized. Nevertheless, we can assess in a standardized manner several features of such organizations that are universal for each stage and show how they differ between stages. The purpose of this paper is to demonstrate that such assessments can be made.

It is still true that in educational evaluation we are most often interested in the overall quantity of knowledge and in the capacity to acquire it. If the overall quantity is high relative to the capacity, then education is deemed "good," if not, then it is deemed "poor."

Traditionally, we have used empirically derived standardized tests to make this comparison between quantity and capacity. Empirical refers to "based on sensation." Organization must be abstracted from sensory content. Therefore, empirically derived evaluation instruments are concerned with the content of knowledge rather than its organization. By standardizing evaluation instruments we make the unit of measurement the same for all users. This facilitates comparisons between users. By testing we make the unit of measurement a correct answer to a question. Quantity then refers to the number of correct answers. If the questions and correct answers are delimited and equated for difficulty for all users then the test is considered standardized. It is generally accepted that if it is possible to test all knowledge of a subject matter then the subject matter is trivial. Empirically derived tests skirt this problem for non-trivial subject matters by sampling from

---

*The authors are indebted to Clark Mitchell and to the faculty of the Educational Research and Demonstration Center for their participation in all phases of this research and to William Derrick for his long-term administrative support of this project.

The realm of possible questions. This procedure creates a host of problems.

The more knowledge a person has of a subject matter the more it is organized. Imagine, for example, a large library that is not organized--it would not be functional. Sampling cannot be done meaningfully unless the organization is known. Imagine pulling cards from the card file of such a library at random and asking a question about each pulled card. The topics of your questions might be something like the following: St. Foix's biography of Mozart, Philosophy, Popular Mechanics for 1957, Iranian Folklore, etc. The procedure becomes only a little less absurd when it is known which subjects 50% of the people know something about. In spite of these problems, we have steadfastly clung to empirically derived standardized tests in the belief that it is not possible to know much about how people's knowledge is organized.

Developmental instruments are based on inferences about changes from simple to more complex organizations. Therefore, developmentally derived evaluation instruments are concerned with the changing organization of knowledge rather than with its content. Developmental theories are either based on children's responses to questions (such as Piaget's and Kohlberg's) or on features of children's natural, spontaneous expressions (such as Lowenfeld's and Moffett's). Therefore, developmentally derived instruments can be standardized in the same way as empirically derived tests (by delimiting possible questions and equating the difficulty of them). Alternatively, they can be standardized by delimiting possible features and determining the stage of these. Since developmental theories specify levels of organization it is possible to choose from them all of a group of general organizational features of non-trivial subjects that cover all or most of the subjects. Therefore, developmentally derived instruments do not require us to choose a small sample from a large realm of possibilities. The conflict between random sampling and levels of organization does not exist. Our contention is first, that if empirically derived tests were preceded by developmental assessments and if the sampling for them

were stratified by what they show about the organization of children's knowledge then such tests would be much more useful. Secondly, even in the absence of such stratified tests, knowledge derived from developmental assessments can be highly useful.

Methods

We began by recognizing that direct empirical derivation of organization is not possible. Thus, our starting point was developmental theories instead of lists of content. Next we made a choice based on experience. In some modes of development there are both of the types of theories mentioned above: those based on responses to questions and those based on children's natural responses. In such cases theories based on children's natural expressions tended to include all the concepts of theories based on experiments and contained many additional concepts. This conclusion was subsequently experimentally tested and supported in an excellent M.A. thesis by Alan Leeds which involved a comparison of the theories of the development of spatial representation by Lowenfeld (expression based) and by Piaget (question based). In addition, no intervention in the classroom, such as presenting a list of questions, was required to obtain natural expressions. Thus, the unit of measurement became the presence of a feature of organization in a child's expression rather than a correct answer to a question. The third step was to develop a procedure to insure that the features assessed were the same for each child--i.e. to standardize the list of features. Finally, the age and stage relations of these features were determined.

Standardized developmental ratings have been constructed by us for two modes of development: drawing and discourse. For each the procedure was logically the same. The primary theories were Lowenfeld's for drawing, and Moffett's for discourse. Each theory can be viewed as consisting of several dimensions, each consisting of several developmentally ordered features. From Lowenfeld's theory

we chose the dimensions of shape, dimensionality, distance, and meaning. From Moffett's we chose the dimensions of speaker-listener distance, speaker-subject distance, speaker-listener-subject distance, logic, rhetoric, and cognition. Because past experience with casual ratings using these theories suggested some important areas of incompleteness, two additional dimensions were added to each theory: composition and design to Lowenfeld's, and meaning and syntactic transformations to Moffett's. An initial positivistic definition for each feature was then written. Following this 10 or 12 expressions from children of different ages were independently assessed for the presence or absence of each feature by a group of 7 to 12 demonstration school faculty. Discrepencies in the ratings were discussed to determine whether they were due to ambiguities in the definitions or to misinterpretations. In the former case, redefinitions were made. The cycle of redefinition-rating-discussion-redefinition was repeated through 5-10 small samples until the ambiguities were eliminated. A large sample of 80-120 expressions (drawings or stories) was then collected from children aged 4-14 and rated by 5 teachers who had been involved in the construction of the list of definitions. These results were analyzed for inter-rater reliability, age-relatedness of features, and stages of development. The analytical procedures have been described elsewhere (Dirlam, et. al., 1976 and 1977).

Results

The results showed that each feature could be assessed with 90% agreement. If two features in the same dimension were not found to differentiate ages, they were collapsed into one. The stages are summarized in the appendix. Examples follow to illustrate these.

Examine the two drawings appended to the paper. The composition in the first drawing shows filled parts but no parts in those parts, while the second drawing has both features. The shapes in the first drawing are made with an overuse of

sample geometric shapes (e.g. the rectangular tree trunk, triangle and rectangle for the house, the circles for traffic lights, etc.) while the second drawing is not. Height on the page only indicates the height of objects in the first drawing and overlap is the only indication of relative distance of objects from the viewer. In the second drawing height indicates distance as well as the height of objects and in addition, distance is shown by size gradient and perspective as well as overlap. The two pictures do not differ in the development of meaning or design. They both have shared meanings but no symbols and neither has decorations.

Now examine the two appended stories. A comparison of them indicates a number of differences in terms of developmental features. On the distance dimensions story 1 is difficult for others to understand, combines present and past events inconsistently, and makes no attempt to corroborate statements with an outside view. Story 2, however, is written for an abstract audience, describes more than one persons view, and occurs in the past or on the "happened" level. On the abstraction dimension, story 1 is disorganized in terms of time and topic, has no definite ending, and is being constructed on an on-going basis. In terms of meaning and syntax, the story is literal with some phrase transformations. Story 2 is organized chronologically and told as a narrative; it maintains the characters and significant objects throughout, and uses metaphors and clause transformations. We are now in the process of constructing training manuals to minimize the misinterpretation of definitions. It appears that practice, feedback and discussion of approximately 30-40 expressions are necessary to obtain reliable ratings by teachers. After such practice it takes a fraction of a minute to rate most drawings and a minute or two to rate most samples of children's discourse.

Discussion

Standardized developmental ratings (SDR's) can be done with a high degree of accuracy. Although we have local age norms, we have not stressed this aspect

of research. For classroom purposes, large discrepencies between the development of individual children and the average age for each stage are noticeable without such norms. Our more immediate goal is to assist teachers in knowing a child's developmental level and in being able to detect when the child is making the transition to the next stage.

Several additional observations can be made concerning the usefulness of SDR's. First, except for the rare case that teachers have already organized their perceptions of children according to developmental theory, learning to SDR's reliably has resulted in a major reorganization of his or her perception of children's drawing and discourse. In addition, most college students spontaneously mention several of our developmental dimensions when being asked to list either what is most important for good writing and drawing, or what is not traditionally reliably evaluated. In contrast, they mention dimensions of achievement testing when asked to list either what can be reliably evaluated or what is less important for good writing. Taken together these findings suggest that SDR's cause teachers to reorganize their perceptions of children's work to focus on intuitively more important aspects of it. Since all the teachers involved in our work have had courses in developmental theories, this observation suggests that SDR's can contribute significantly to the internalization and use of such theories.

Secondly, a major practical use of SDR's can be derived from the well-known experimental finding that developing organisms are most vulnerable during times of rapid developmental transitions. This would suggest that the time to push children to elaborate their knowledge is within stages and not between them and that it is highly inappropriate to push children to the next stage before they are ready.

Thirdly, a series of findings have significant implications for the use of creative processes in education. One of the teachers involved in this project (Rhoda Mead) has conducted an experiment showing that children are more able to do assign-

ments at or even above their developmental level than below. When asked to do assignments appropriate to younger children, older children expect them to be easy, are unable to do them, and are unaware that they haven't done them. In contrast, Clark Mitchell and I have found that adults can be induced to function on lower developmental levels by being placed in situations demanding much detailed knowledge in very short time periods. Finally, in an assessment of the paintings of Matisse, Miro, Hopper, Picasso, and Chagall, it is notable that all of them mix developmental levels. These findings support the temporary regression theory of creativity and suggest that such regression can be induced but not by direct instruction.

Fourthly, a current limitation of this research is the age range restriction to 4-14 years. Further work is needed with both older and younger ages before the stages should be considered definitive.

In conclusion, children naturally discourse and draw. Doing SDR's requires no intervention but gives standardized results on features that teachers consider very important. They appear to have a profound effect on teachers' concepts of their practice but as yet research on how they alter that practice has only just begun. One important area that is still unexplored is the effect of school reporting of such assessments on parents' attitudes toward their children. Finally, SDR's can be used to provide quantitative summaries for evaluators and policy makers, but they appear to be far more important for their ability to provide precise and objective qualitative information.

References

Dirlam, D., Byrne M., Mitchell, J. C., Clark, S., and Mead, R. Development Based Education: II. Discourse and its relation to art. Symposium presented at Northeastern Educational Research Association, October, 1977.

Dirlam, D., Mitchell, J. C., and Olson, H. Development Based Education: I. Spatial Representation. Symposium presented at Northeastern Educational Research Association, October 1976.

Lowenfeld, V. Creative and Mental Growth, Third Edition. New York: MacMillan, 1957.

Moffett, J. Teaching the Universe of Discourse. Boston: Houghton Mifflin, 1968.

Piaget, J. and Inhelder, B. The child's conception of space. (Transl. by F. S. Langdon & J. L. Lurrzer). London: Routledge & Kegan Paul, 1956 (Originally published in 1948).

APPENDIX:

Two drawings by children

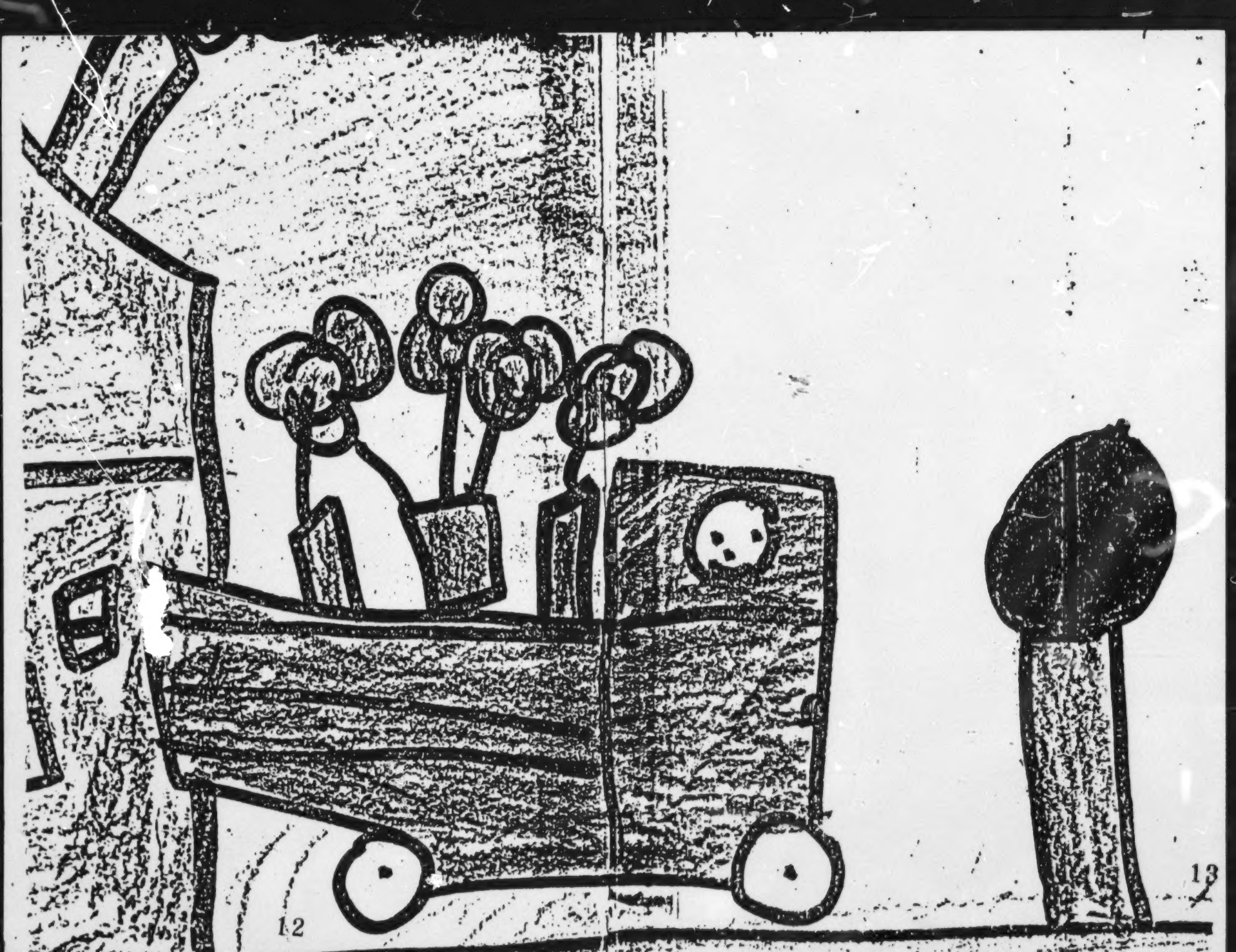
Two stories by children

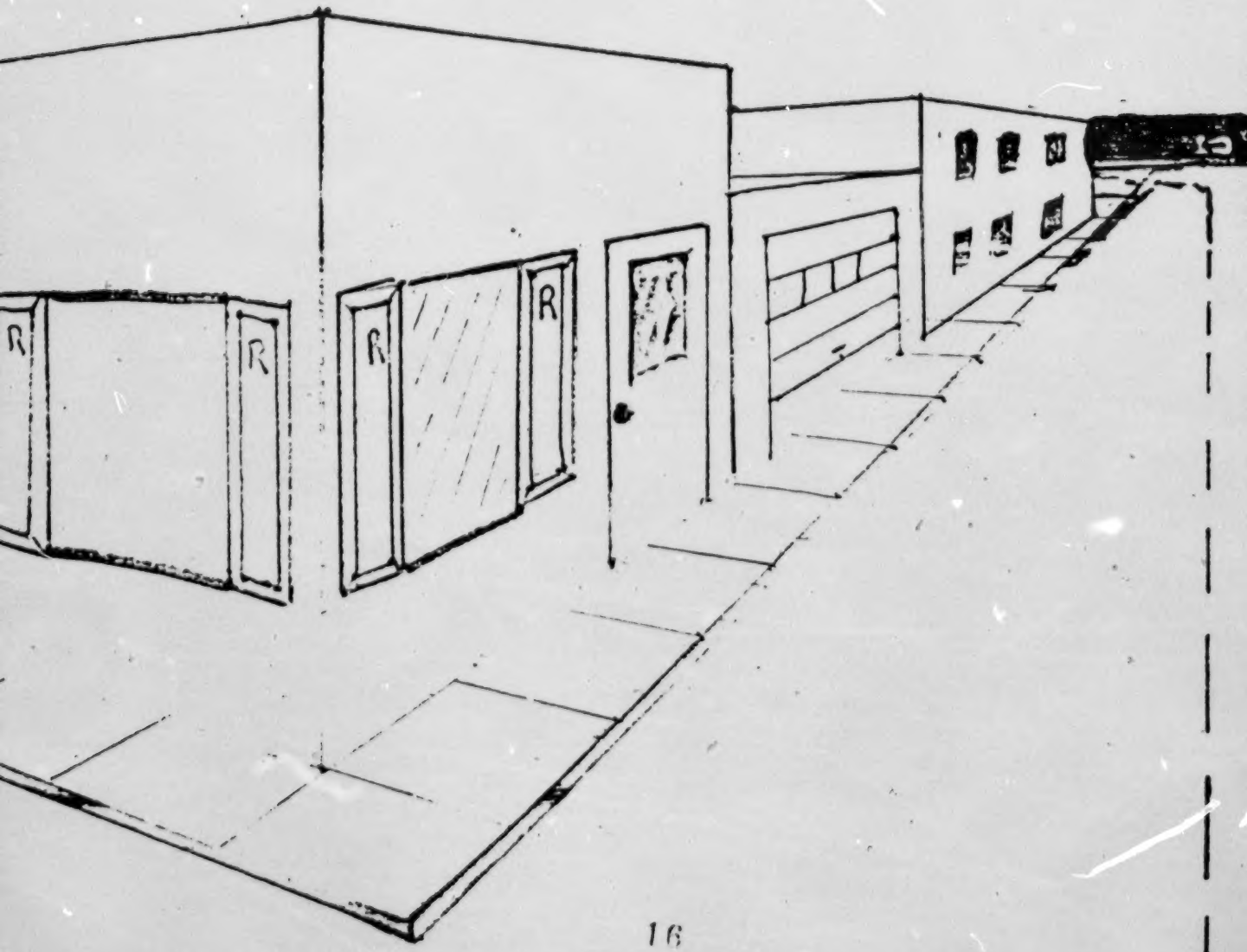
Summary of Stages in the development of drawing

Summary of Stages in the development of discourse

Definitions of features of the development of drawing

Definitions of features of the development of discourse




R
R
R
R

Story 1

There was a coconut and a cuckoo tree in our front yard and there were that many cuckoo birds (holding up three fingers) and I shake the tree and the coconuts came coming down falling on my head and on my nose. And there were other birds that didn't make a nest in the tree--they made a nest on the ground. That's the thing that fell on my head (pointing to the coconut)--they (meaning the birds) thought it was a bomb but it wasn't they thought they could make a little hole in it. They made three holes in it. The juice is in here (picking up the coconut) that's their food.

A few minutes passed and Ben said, "I want to tell you something about that." "That coconut is rich--I got it from Florida."

Story 2

"The Ice Mountain Princess"

Once there was an enchanted princess who lived on an ice hill. It was much too slippery for anyone to climb. So the girl's suitors would have to sit at the foot of the mountain and call up their praise. Each summer the ground on the hill thawed and water flooded the village. After many years of flood the village decided to destroy the mountain. 11,000 men with 11,000 hammers, and 30,000 oxen poured into the city and went about shattering the glass. Young women sat on the side of the road with envy in their heart and young boys awaited their beauty to come down from the hill.

All at once the hill shattered and the castle at the top fell upon the 30,000 oxen and 11,000 men with 11,000 shovels. The princess stepped out of the rubble with shiny tears making runs on her crystal-like cheeks. She was clothed in a dress of snake silk and a veil of spider's web. her shoes were woven lace and her blonde hair fell down over her shoulders. The men rushed forward but she only passed them by and had no vengence feelings towards the snickering women.

A Brief Outline of the Development of Drawing

Educational Research and Demonstration Center
State University College of Arts and Science

Stage Ia is characterized by either scribbles or simple closed objects with no parts. At this stage there are no dimensions, no indication of distance and no decoration. Drawings may or may not be such that any viewer would know the meaning of them.

By Stage Ib objects will have geometric parts and shared meanings. Objects will either be filled or will be located on baselines but they will not be both filled and located on baselines. There still is no indication of distance and no decoration.

Stage IIa drawings have both filled parts and baselines. There may be realistic outlines as well as geometric shapes; color may, but is not necessarily, used meaningfully; and there may be decoration of objects. There is still no indication of distance.

In Stage IIb, distance of objects is indicated, in representational drawings, by overlap. The lines of one object are broken by the outline of another (if one object is just drawn over another it is an x-ray, not an overlapping, drawing). By this stage objects may lose their overdependence on geometric shapes and base planes may replace the baseline. All use of color will be meaningful and there may be decoration. In some cases decorative designs without symmetry or meaningful objects may be made.

In Stage IIc all shapes will be realistic and all representational drawings will have base planes. Objects will only be constructed of triangles, rectangles and circles only if they appear so in the natural world. Distance will be indicated by the height of an object as well as by overlap and there may even be a size gradient (objects intended to be more distant are drawn smaller). Also joints may be shown to indicate action of objects that move in the natural world. Decorative designs

A Brief Outline of the Development of Discourse

Educational Research and Demonstration Center
State University College of Arts and Science

Stage I is characterized primarily as discourse directed toward the self, although some stories are directed toward others. The story is being constructed on an on-going basis, makes no attempt to corroborate statements with an outside view, and frequently is disorganized in terms of time and/or topic. Frequently, there is no definite ending to the story. Individual words may be used either preliterally or literally, with some word transformations present.

Stage II directs the discourse to a specific audience, compares the views of two or more participants within the discourse about the same events, and arranges events in a chronological or etiological order. Both subject matter and time sequence are retained throughout. Individual words are used on a literal level and phrase transformations are evident.

Stage III discourse is written about one or several other(s) to an abstract or unknown audience, depends on past events, is done from memory or on generalizations constructed before the discourse, and may include interpretations of experiences common to large segments of humanity. There is a clear beginning and ending within a narrative or fable format. Individual word usage is literal and clause transformations are included.

Stage IV discourse presents an abstract subject to an abstract audience, describes and interprets possibilities, functions on a conjectural level, and explores universal human experiences. Abstract concepts are related to concrete events, such as with an outline or to other abstract concepts, as in a deductive sense. The implications of a group of generalizations are examined. Metaphors or analogies are present and 3 or more clauses arranged logically in terms of importance can be identified.